



-118-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Ponath, Paul D.
Ringler, Douglas J.
Jones, S. Tarran
Newman, Walter
Saldanha, José
Bendig, Mary M.

(ii) TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH $\alpha 4\beta 7$
INTEGRIN

(iii) NUMBER OF SEQUENCES: 63

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
(B) STREET: Two Militia Drive
(C) CITY: Lexington
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02173

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/700,737
(B) FILING DATE: 15-AUG-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Brook, David E.
(B) REGISTRATION NUMBER: 22,592
(C) REFERENCE/DOCKET NUMBER: LKS95-10

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 861-6240
(B) TELEFAX: (617) 861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 494 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 13..444

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

(xi) SEQUENCE DESCRIPTION: SEE LIST							
TTACKRGWMK	WCATGRRATG	SASCTRKRTC	ATYYTCTTCT	TGGTATCAAC	AGCTACAAGT		60
GTCCAATCCC	AGGTCCAAC	GCAGCAGCCT	GGGGCTGAGC	TTGTGAAGCC	TGGGACTTCA		120
GTGAAGCTGT	CCTGCAAGGG	TTATGGCTAC	ACCTTCACCA	GCTACTGGAT	GCACTGGGTG		180
AAGCAGAGGC	CTGGACAAGG	CCTTGAGTGG	ATCGGAGAGA	TTGATCCTTC	TGAGAGTAAT		240
ACTAACTACA	ATCAAAAATT	CAAGGGCAAG	GCCACATTGA	CTGTAGACAT	TTCCTCCAGC		300
ACAGCCTACA	TGCAGCTCAG	CAGCCTGACA	TCTGAGGACT	CTGCGGTCTA	CTATTGTGCA		360
AGAGGGGGTT	ACGACGGATG	GGACTATGCT	ATTGACTACT	GGGGTCAAGG	CACCTCAGTC		420
ACCGTCTCCT	CAGCCAAAAC	GACACCRYCN	CSYKTMTCYC	YYSBDNNCCC	YKGRWSCYTG		480
GNNGAAGCTT	GGGA						494

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- SEQUENCE CHARACTERISTICS.
- (A) LENGTH: 144 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

(xi) SEQUENCE DESCRIPTION: 112

Met Xaa Xaa Xaa Xaa Xaa Ile Xaa Phe Leu Val Ser Thr Ala Thr Ser
1 5 10 15

Val His Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys
20 25 30

Pro Gly Thr Ser Val Lys Leu Ser Cys Lys Gly Tyr Gly Tyr Thr Phe
35 40 45

Thr Ser Tyr Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu
50 55 60

Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn
65 70 75 80

Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Ile Ser Ser Ser
85 90 95

Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val
100 105 110

Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp
115 120 125

Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr Thr
130 135 140

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 428 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 18..428

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTACTTGACG ACTCGGG ATG GGA TGG AGC TAT ATC ATC TTC TTC TTG GTA	50
Met Gly Trp Ser Tyr Ile Ile Phe Phe Leu Val	10
1	
TCA ACA GCT ACA AGT GTC CAC TCC CAG GTC CAA CTG CAG CAG CCT GGG	98
Ser Thr Ala Thr Ser Val His Ser Gln Val Gln Leu Gln Gln Pro Gly	25
15	
GCT GAG CTT GTG AAG CCT GGG ACT TCA GTG AAG CTG TCC TGC AAG GGT	146
Ala Glu Leu Val Lys Pro Gly Thr Ser Val Lys Leu Ser Cys Lys Gly	40
30	
TAT GGC TAC ACC TTC ACC AGC TAC TGG ATG CAC TGG GTG AAG CAG AGG	194
Tyr Gly Tyr Thr Phe Thr Ser Tyr Trp Met His Trp Val Lys Gln Arg	55
45	
CCT GGA CAA GGC CTT GAG TGG ATC GGA GAG ATT GAT CCT TCT GAG AGT	242
Pro Gly Gln Gly Leu Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser	75
60	
AAT ACT AAC TAC AAT CAA AAA TTC AAG GGC AAG GCC ACA TTG ACT GTA	290
Asn Thr Asn Tyr Asn Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val	90
80	
GAC ATT TCC TCC AGC ACA GCC TAC ATG CAG CTC AGC AGC CTG ACA TCT	338
Asp Ile Ser Ser Ser Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser	105
95	
GAG GAC TCT GCG GTC TAC TAT TGT GCA AGA GGG GGT TAC GAC GGA TGG	386
Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp	120
110	
GAC TAT GCT ATT GAC TAC TGG GGT CAA GGC ACA TCA GTC ACC	428
Asp Tyr Ala Ile Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr	135
125	

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 137 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Trp Ser Tyr Ile Ile Phe Phe Leu Val Ser Thr Ala Thr Ser
 1 5 10 15
 Val His Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys
 20 25 30
 Pro Gly Thr Ser Val Lys Leu Ser Cys Lys Gly Tyr Gly Tyr Thr Phe
 35 40 45
 Thr Ser Tyr Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu
 50 55 60
 Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn
 65 70 75 80
 Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Ile Ser Ser Ser
 85 90 95
 Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val
 100 105 110
 Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp
 115 120 125
 Tyr Trp Gly Gln Gly Thr Ser Val Thr
 130 135

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 535 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 16..435

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGATTACTAG TCGAC ATG AAG TTG CCT GTT AGG CTG TTG GTG CTT CTG TTG 51
 Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Leu 145
 TTC TGG ATT CCT GTT TCC GGA GGT GAT GTT GTG GTG ACT CAA ACT CCA 99
 Phe Trp Ile Pro Val Ser Gly Gly Asp Val Val Val Thr Gln Thr Pro 165
 150 155 160
 CTC TCC CTG CCT GTC AGC TTT GGA GAT CAA GTT TCT ATC TCT TGC AGG 147
 Leu Ser Leu Pro Val Ser Phe Gly Asp Gln Val Ser Ile Ser Cys Arg 180
 170 175
 TCT AGT CAG AGT CTT GCA AAG AGT TAT GGG AAC ACC TAT TTG TCT TGG 195
 Ser Ser Gln Ser Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp 195
 185 190

TAC CTG CAC AAG CCT GGC CAG TCT CCA CAG CTC CTC ATC TAT GGG ATT	243
Tyr Leu His Lys Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile	
200 205 210	
TCC AAC AGA TTT TCT GGG GTG CCA GAC AGG TTC AGT GGC AGT GGT TCA	291
Ser Asn Arg Phe Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser	
215 220 225	
GGG ACA GAT TTC ACA CTC AAG ATC AGC ACA ATA AAG CCT GAG GAC TTG	339
Gly Thr Asp Phe Thr Leu Lys Ile Ser Thr Ile Lys Pro Glu Asp Leu	
230 235 240 245	
GGA ATG TAT TAC TGC TTA CAA GGT ACA CAT CAG CCG TAC ACG TTC GGA	387
Gly Met Tyr Tyr Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly	
250 255 260	
GGG GGG ACC AAG CTG GAA ATA AAA CGG GCT GAT GCT GCA CCA ACT GTA	435
Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp Ala Ala Pro Thr Val	
265 270 275	
TCCATCTTCC CACCATCCAG TAAGCTTGGG AATCCATATG ACTAGTAGAT CCTCTAGAGT	495
CGACCTGCAG GCATGCAAGC TTCCCTATAG TGAGTCGTAT	535

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 140 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Leu Phe Trp Ile Pro	15
1 5 10	
Val Ser Gly Gly Asp Val Val Val Thr Gln Thr Pro Leu Ser Leu Pro	30
20 25	
Val Ser Phe Gly Asp Gln Val Ser Ile Ser Cys Arg Ser Ser Gln Ser	45
35 40	
Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu His Lys	60
50 55	
Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe	80
65 70 75	
Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe	95
85 90	
Thr Leu Lys Ile Ser Thr Ile Lys Pro Glu Asp Leu Gly Met Tyr Tyr	110
100 105	
Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly Gly Gly Thr Lys	125
115 120	
Leu Glu Ile Lys Arg Ala Asp Ala Ala Pro Thr Val	140
130 135	

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 112 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Asp Val Val Val Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Phe Gly
1           5           10           15
Asp Gln Val Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Ala Lys Ser
20           25           30
Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu His Lys Pro Gly Gln Ser
35           40           45
Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe Ser Gly Val Pro
50           55           60
Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65           70           75           80
Ser Thr Ile Lys Pro Glu Asp Leu Gly Met Tyr Tyr Cys Leu Gln Gly
85           90           95
Thr His Gln Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100          105          110
  
```

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 112 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Asp Ile Val Met Thr Gln Ser Pro Leu Ser Leu Pro Val Thr Pro Gly
1           5           10           15
Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
20           25           30
Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln Ser
35           40           45
Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg Ala Ser Gly Val Pro
50           55           60
  
```

Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65 70 75 80
Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Met Gln Ala
85 90 95
Leu Gln Thr Pro Gln Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
100 105 110

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys Pro Gly Thr
1 5 10 15
Ser Val Lys Leu Ser Cys Lys Gly Tyr Gly Tyr Thr Phe Thr Ser Tyr
20 25 30
Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile
35 40 45
Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn Gln Lys Phe
50 55 60
Lys Gly Lys Ala Thr Leu Thr Val Asp Ile Ser Ser Ser Thr Ala Tyr
65 70 75 80
Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp Tyr Trp Gly
100 105 110
Gln Gly Thr Ser Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
20 25 30
Ala Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Met
35 40 45
Gly Trp Ile Asn Ala Gly Asn Gly Asn Thr Lys Tyr Ser Gln Lys Phe
50 55 60
Gln Gly Arg Val Thr Ile Thr Arg Asp Thr Ser Ala Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Gly Gly Tyr Tyr Gly Ser Gly Ser Asn Tyr Trp Gly Gln Gly
100 105 110
Thr Leu Val Thr Val Ser Ser
115

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 396 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..396

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG AAG TTG CCT GTT AGG CTG TTG GTG CTT CTG TTG TTC TGG ATT CCT	48
Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Leu Phe Trp Ile Pro	155
GTT TCC GGA GGT GAT GTT GTG GTG ACT CAA ACT CCA CTC TCC CTG CCT	96
Val Ser Gly Gly Asp Val Val Val Thr Gln Thr Pro Leu Ser Leu Pro	170
GTC AGC TTT GGA GAT CAA GTT TCT ATC TCT TGC AGG TCT AGT CAG AGT	144
Val Ser Phe Gly Asp Gln Val Ser Ile Ser Cys Arg Ser Ser Gln Ser	185
CTT GCA AAG AGT TAT GGG AAC ACC TAT TTG TCT TGG TAC CTG CAC AAG	192
Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu His Lys	200
CCT GGC CAG TCT CCA CAG CTC CTC ATC TAT GGG ATT TCC AAC AGA TTT	240
Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe	220
205 210 215	

TCT GGG GTG CCA GAC AGG TTC AGT GGC AGT GGT TCA GGG ACA GAT TTC	288
Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe	
225 230 235	
ACA CTC AAG ATC AGC ACA ATA AAG CCT GAG GAC TTG GGA ATG TAT TAC	336
Thr Leu Lys Ile Ser Thr Ile Lys Pro Glu Asp Leu Gly Met Tyr Tyr	
240 245 250	
TGC TTA CAA GGT ACA CAT CAG CCG TAC ACG TTC GGA GGG GGG ACC AAG	384
Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly Gly Gly Thr Lys	
255 260 265	
CTG GAA ATA AAA	396
Leu Glu Ile Lys	
270	

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 132 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Phe Trp Ile Pro	15
1 5 10	
Val Ser Gly Gly Asp Val Val Val Thr Gln Thr Pro Leu Ser Leu Pro	30
20 25 30	
Val Ser Phe Gly Asp Gln Val Ser Ile Ser Cys Arg Ser Ser Gln Ser	45
35 40 45	
Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu His Lys	60
50 55 60	
Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe	80
65 70 75 80	
Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe	95
85 90 95	
Thr Leu Lys Ile Ser Thr Ile Lys Pro Glu Asp Leu Gly Met Tyr Tyr	110
100 105 110	
Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly Gly Gly Thr Lys	125
115 120 125	
Leu Glu Ile Lys	
130	

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 336 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GATATTGTGA TGA	CTCAGTC	TCCACTCTCC	CTGCCCCGTCA	CCCCTGGAGA	GCCGGCCTCC	60
ATCTCCTGCA GGT	CTAGTCA	GAGCCTCCTC	CATAGTAATG	GATCAAAC	TTTGGATTGG	120
TACCTGCAGA AGC	CAGGGCA	GTCTCCACAG	CTCCTGATCT	ATTTGGGTTC	TAATCGGGCC	180
TCCGGGGTCC CTG	ACAGGTT	CAGTGGCAGT	GGATCAGGCA	CAGATTTTAC	ACTGAAAATC	240
AGCAGAGTGG AGG	CTGAGGA	TGTTGGGGTT	TATTACTGCA	TGCAAGCTCT	ACCAACTCCT	300
CAGACGTTTCG	GCCAAGGGAC	CAAGGTGGAA	ATCAAA			336

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 420 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..420

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATG GGA TGG AGC TGT ATC ATC CTC TTC TTG GTA TCA ACA GCT ACA AGT	48
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ser Thr Ala Thr Ser	
135 140 145	
GTC CAC TCC CAG GTC CAA CTG CAG CAG CCT GGG GCT GAG CTT GTG AAG	96
Val His Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys	
150 155 160	
CCT GGG ACT TCA GTG AAG CTG TCC TGC AAG GGT TAT GGC TAC ACC TTC	144
Pro Gly Thr Ser Val Lys Leu Ser Cys Lys Gly Tyr Gly Tyr Thr Phe	
165 170 175 180	
ACC AGC TAC TGG ATG CAC TGG GTG AAG CAG AGG CCT GGA CAA GGC CTT	192
Thr Ser Tyr Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu	
185 190 195	
GAG TGG ATC GGA GAG ATT GAT CCT TCT GAG AGT AAT ACT AAC TAC AAT	240
Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn	
200 205 210	
CAA AAA TTC AAG GGC AAG GCC ACA TTG ACT GTA GAC ATT TCC TCC AGC	288
Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Ile Ser Ser Ser	
215 220 225	
ACA GCC TAC ATG CAG CTC AGC AGC CTG ACA TCT GAG GAC TCT GCG GTC	336
Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val	
230 235 240	

TAC TAT TGT GCA AGA GGG GGT TAC GAC GGA TGG GAC TAT GCT ATT GAC 384
Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp 260
245 250 255

TAC TGG GGT CAA GGC ACC TCA GTC ACC GTC TCC TCA 420
Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser 270
265

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ser Thr Ala Thr Ser
1 5 10 15
Val His Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys
20 25 30
Pro Gly Thr Ser Val Lys Leu Ser Cys Lys Gly Tyr Gly Tyr Thr Phe
35 40 45
Thr Ser Tyr Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu
50 55 60
Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn
65 70 75 80
Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Ile Ser Ser Ser
85 90 95
Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val
100 105 110
Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp
115 120 125
Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
130 135 140

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..414

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATG GAG TTT GGG CTG AGC TGG CTT TTT CTT GTG GCT ATT TTA AAA GGT	48
Met Glu Phe Gly Leu Ser Trp Leu Phe Leu Val Ala Ile Leu Lys Gly	
145 150 155	
GTC CAG TGT CAG GTG CAG CTT GTG CAG TCT GGG GCT GAG GTG AAG AAG	96
Val Gln Cys Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys	
160 165 170	
CCT GGG GCC TCA GTG AAG GTT TCC TGC AAG GCT TCT GGA TAC ACC TTC	144
Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe	
175 180 185	
ACT AGC TAT GCT ATG CAT TGG GTG CGC CAG GCC CCC GGA CAA AGG CTT	192
Thr Ser Tyr Ala Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu	
190 195 200	
GAG TGG ATG GGA TGG ATC AAC GCT GGC AAT GGT AAC ACA AAA TAT TCA	240
Glu Trp Met Gly Trp Ile Asn Ala Gly Asn Gly Asn Thr Lys Tyr Ser	
205 210 215 220	
CAG AAG TTC CAG GGC AGA GTC ACC ATT ACC AGG GAC ACA TCC GCG AGC	288
Gln Lys Phe Gln Gly Arg Val Thr Ile Thr Arg Asp Thr Ser Ala Ser	
225 230 235	
ACA GCC TAC ATG GAG CTG AGC AGC CTG AGA TCT GAA GAC ACG GCT GTG	336
Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val	
240 245 250	
TAT TAC TGT GCG AGA GGA GGT TAC TAT GGT TCG GGG AGC AAC TAC TGG	384
Tyr Tyr Cys Ala Arg Gly Gly Tyr Tyr Gly Ser Gly Ser Asn Tyr Trp	
255 260 265	
GGC CAG GGA ACC CTG GTC ACC GTC TCC TCA	414
Gly Gln Gly Thr Leu Val Thr Val Ser Ser	
270 275	

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 138 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Glu Phe Gly Leu Ser Trp Leu Phe Leu Val Ala Ile Leu Lys Gly	
1 5 10 15	
Val Gln Cys Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys	
20 25 30	
Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe	
35 40 45	
Thr Ser Tyr Ala Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu	
50 55 60	

Glu Trp Met Gly Trp Ile Asn Ala Gly Asn Gly Asn Thr Lys Tyr Ser
65 70 75 80
Gln Lys Phe Gln Gly Arg Val Thr Ile Thr Arg Asp Thr Ser Ala Ser
85 90 95
Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val
100 105 110
Tyr Tyr Cys Ala Arg Gly Gly Tyr Tyr Gly Ser Gly Ser Asn Tyr Trp
115 120 125
Gly Gln Gly Thr Leu Val Thr Val Ser Ser
130 135

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..540

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATG AAA TGC ACC TGG GTC ATT CTC TTC TTG GTA TCA ACA GCT ACA AGT	48
Met Lys Cys Thr Trp Val Ile Leu Phe Leu Val Ser Thr Ala Thr Ser	
140 145 150	
GTC CAC TCC CAG GTC CAA CTA GTG CAG TCT GGG GCT GAG GTT AAG AAG	96
Val His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys	
155 160 165 170	
CCT GGG GCT TCA GTG AAG GTG TCC TGC AAG GGT TCT GGC TAC ACC TTC	144
Pro Gly Ala Ser Val Lys Val Ser Cys Lys Gly Ser Gly Tyr Thr Phe	
175 180 185	
ACC AGC TAC TGG ATG CAT TGG GTG AGG CAG GCG CCT GGC CAA CGT CTA	192
Thr Ser Tyr Trp Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu	
190 195 200	
GAG TGG ATC GGA GAG ATT GAT CCT TCT GAG AGT AAT ACT AAC TAC AAT	240
Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn	
205 210 215	
CAA AAA TTC AAG GGA CGC GTC ACA TTG ACT GTA GAC ATT TCC GCT AGC	288
Gln Lys Phe Lys Gly Arg Val Thr Leu Thr Val Asp Ile Ser Ala Ser	
220 225 230	
ACA GCC TAC ATG GAG CTC AGC AGC CTG AGA TCT GAG GAC ACT GCG GTC	336
Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val	
235 240 245 250	

TAC TAT TGT GCA AGA GGG GGT TAC GAC GGA TGG GAC TAT GCT ATT GAC Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp 255 260 265	384
TAC TGG GGT CAA GGC ACC CTG GTC ACC GTC TCC TCA GCC TCC ACC AAG Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys 270 275 280	432
GGC CCA TCG GTC TTC CCC CTG GCA CCC TCC TCC AAG AGC ACC TCT GGG Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly 285 290 295	480
GGC ACA GCG GCC CTG GGC TGC CTG GTC AAG GAC TAC TTC CCC GAA CCG Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro 300 305 310	528
GTG ACG GTG TCG Val Thr Val Ser 315	540

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 180 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Lys Cys Thr Trp Val Ile Leu Phe Leu Val Ser Thr Ala Thr Ser
1 5 10 15
Val His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys
20 25 30
Pro Gly Ala Ser Val Lys Val Ser Cys Lys Gly Ser Gly Tyr Thr Phe
35 40 45
Thr Ser Tyr Trp Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu
50 55 60
Glu Trp Ile Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn
65 70 75 80
Gln Lys Phe Lys Gly Arg Val Thr Leu Thr Val Asp Ile Ser Ala Ser
85 90 95
Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val
100 105 110
Tyr Tyr Cys Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp
115 120 125
Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys
130 135 140
Gly Pro Ser Val Phe Pro Leu Ala Pro S r Ser Lys Ser Thr Ser Gly
145 150 155 160

Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro
165 170 175

Val Thr Val Ser
180

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 413 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..413

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATG AAG TTG CCT GTT AGG CTG TTG GTG CTT CTG TTG TTC TGG ATT CCT	48
Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Leu Phe Trp Ile Pro	195
185	
GTT TCC GGA GGT GAT GTT GTG ATG ACT CAA AGT CCA CTC TCC CTG CCT	96
Val Ser Gly Gly Asp Val Val Met Thr Gln Ser Pro Leu Ser Leu Pro	210
200	
GTC ACC CCT GGA GAA CCA GCT TCT ATC TCT TGC AGG TCT AGT CAG AGT	144
Val Thr Pro Gly Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser	225
215	
CTT GCA AAG AGT TAT GGG AAC ACC TAT TTG TCT TGG TAC CTG CAG AAG	192
Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu Gln Lys	240
230	
CCT GGC CAG TCT CCA CAG CTC CTC ATC TAT GGG ATT TCC AAC AGA TTT	240
Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe	260
245	
TCT GGG GTG CCA GAC AGG TTC AGT GGC AGT GGT TCA GGG ACA GAT TTC	288
Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe	275
265	
ACA CTC AAG ATC TCG CGA GTA GAG GCT GAG GAC GTG GGA GTG TAT TAC	336
Thr Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr	290
280	
TGC TTA CAA GGT ACA CAT CAG CCG TAC ACG TTC GGA CAG GGG ACC AAG	384
Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly Gln Gly Thr Lys	305
295	
GTG GAA ATA AAA CGG GCT GAT GCG GCG CC	413
Val Glu Ile Lys Arg Ala Asp Ala Ala Pro	315
310	

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 138 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

Met Lys Leu Pro Val Arg Leu Leu Val Leu Leu Leu Phe Trp Ile Pro
 1           5           10           15
Val Ser Gly Gly Asp Val Val Met Thr Gln Ser Pro Leu Ser Leu Pro
          20           25           30
Val Thr Pro Gly Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser
          35           40           45
Leu Ala Lys Ser Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu Gln Lys
          50           55           60
Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe
          65           70           75           80
Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
          85           90           95
Thr Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr
          100          105          110
Cys Leu Gln Gly Thr His Gln Pro Tyr Thr Phe Gly Gln Gly Thr Lys
          115          120          125
Val Glu Ile Lys Arg Ala Asp Ala Ala Pro
          130          135
  
```

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 94 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

TTTCCGGAGG TGATGTTGTG ATGACTCAAA GTCCACTCTC CCTGCCTGTC ACCCCTGGAG
AACCAGCTTC TATCTCTTGC AGGTCTAGTC AGAG
  
```

60

94

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 94 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23: 60
ACTGGCCAGG CTTCTGCAGG TACCAAGACA AATAGGTGTT CCCATAACTC TTTGCAAGAC 94
TCTGACTAGA CCTGCAAGAG ATAGAAGCTG GTTC

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 83 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24: 60
CCTGGCCAGT CTCCACAGCT CCTCATCTAT GGGATTTC CAACAGATTTTC TGGGGTGCCA 83
GACAGGTTCA GTGGCAGTGG TTC

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 84 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25: 60
ACTCGCGAGA TCTTGAGTGT GAAATCTGTC CCTGAACCAC TGCCACTGAA CCTGTCTGGC 84
ACCCAGAAAA ATCTGTTGGA AATC

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26: 60
TCTCGCGAGT AGAGGCTGAG GACGTGGGAG TGTATTACTG CTTACAAGGT ACACATCAGC 67
CGTACAC

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 86 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGGCGCCGC ATCAGCCCGT TTTATTTCCA CCTTGGTCCC CTGTCCGAAC GTGTACGGCT 60
GATGTGTACC TTGTAAGCAG TAATAC 86

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 93 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATAAGCTTCG CCATGAAATG CACCTGGGTC ATTCTCTTCT TGGTATCAAC AGCTACAAGT 60
TGCCACTCCC AGGTCCAAC AGTGCACCGG TTA 93

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 93 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TAACCGGTGC ACTAGTTGGA CCTGGGAGTG GACACTTGTA GCTGTTGATA CCAAGAAGAG 60
AATGACCCAG GTGCATTTC TGGCGAAGCT TAT 93

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 87 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30: 60
CAACTAGTGC AGTCTGGGGC TGAGGTTAAG AAGCCTGGGG CTCAGTGAA GGTGTCCTGC 87
AAGGGTTCTG GCTACACCTT CACCAGC

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31: 60
TAACCGGTAC TCTAGACGTT GGCCAGGCGC CTGCCTCACC CAATGCATCC AGTAGCTGGT 88
GAAGGTGTAG CCAGAACCCT TGCAGGAC

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32: 60
CGTCTAGAGT GGATCGGAGA GATTGATCCT TCTGAGAGTA ATACTAACTA CAATCAAAAA 76
TTCAAGGGAC GCGTCA

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33: 60
TAACCGGTGT GCTAGCGGAA ATGTCTACAG TCAATGTGAC GCGTCCCTTG AATTTTGTAT 87
TG TAGTTAGT ATTACT

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 88 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CCGCTAGCAC AGCCTACATG GAGCTCAGCA GCCTGAGATC TGAGGACACT GCGGTCTACT 60
ATTGTGCAAG AGGGGGTTAC GACGGATG 88

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 88 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TCACCGGTGC GGTGACCAGG GTGCCTTGAC CCCAGTAGTC AATAGCATAG TCCCATCCGT 60
CGTAACCCCC TCTTGACAA TAGTAGAC 88

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 85 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CTGGTCACCG TCTCCTCAGC CTCCACCAAG GGCCATCGG TCTCCCCCT GGCACCTCC 60
TCCAAGAGCA CCTCTGGGGG CACAG 85

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 85 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37: 60
TCACCGGTTTC GGGGAAGTAG TCCTTGACCA GGCAGCCCAG GGCCGCTGTG CCCCCAGAGG 85
TGCTCTTGGA GGAGGGTGCC AGGGG

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38: 10
CTGGCCAACG

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39: 34
CACATTGACT GTAGACACTT CCGCTAGCAC AGCC

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40: 24
CCGGAGGTGA TGTTGTGGTG ACTC

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TAAGCTTCCG CCATGGGATG GAGC

24

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GGTGACACTA GTGCCTTGAC CCCAG

25

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TAAGCTTCCG CCATGAAGTT GCCT

24

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGCCGCAT CAGCCCGTTT T

21

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CGGCGCCATC TGTCTTCATC

20

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AAGCTTCTAA CACTCTCC

18

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Asp	Val	Val	Val	Thr	Gln	Thr	Pro	Leu	Ser	Leu	Pro	Val	Ser	Phe	Asp
1				5				10						15	

Gly Gln Val

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Asp	Val	Val	Val	Thr	Gln	Thr	Pro	L	u	Ser	Leu
1				5						10	

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Asp Tyr Ala Ile Asp Tyr Trp Gly
 1 5

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 113 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Asp Val Val Met Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Leu Gly
 1 5 10 15
 Asp Gln Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Val His Ser
 20 25 30
 Asn Gly Asn Thr Tyr Leu Glu Trp Tyr Leu Gln Lys Pro Gly Gln Ser
 35 40 45
 Pro Lys Leu Leu Ile Tyr Lys Val Ser Asn Arg Phe Ser Gly Val Pro
 50 55 60
 Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
 65 70 75 80
 Ser Arg Val Glu Ala Glu Asp Leu Gly Val Tyr Tyr Cys Phe Gln Gly
 85 90 95
 Thr His Val Pro Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
 100 105 110
 Lys

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 114 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Asp Ile Val Met Thr Gln Ser Pro Leu Ser Leu Pro Val Thr Pro Gly
 1 5 10 15
 Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
 20 25 30
 Xaa Asp Gly Asn Asn Tyr Leu Asn Trp Tyr Leu Gln Lys Pro Gly Gln
 35 40 45
 Ser Pro Gln Leu Leu Ile Tyr Leu Val Ser Asn Arg Ala Ser Gly Val
 50 55 60
 Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys
 65 70 75 80
 Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Met Gln
 85 90 95
 Ala Leu Gln Xaa Pro Arg Xaa Thr Phe Gly Gln Gly Thr Lys Val Glu
 100 105 110
 Ile Lys

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 112 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Asp Val Val Met Thr Gln Ser Pro Leu Ser Leu Pro Val Thr Pro Gly
 1 5 10 15
 Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Ala Lys Ser
 20 25 30
 Tyr Gly Asn Thr Tyr Leu Ser Trp Tyr Leu Gln Lys Pro Gly Gln Ser
 35 40 45
 Pro Gln Leu Leu Ile Tyr Gly Ile Ser Asn Arg Phe Ser Gly Val Pro
 50 55 60
 Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
 65 70 75 80
 Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Leu Gln Gly
 85 90 95

Thr His Gln Pro Tyr Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
100 105 110

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 127 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys Pro Gly Ala
1 5 10 15
 Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
20 25 30
 Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile
35 40 45
 Gly Arg Ile Asp Pro Asn Ser Gly Gly Thr Asn Tyr Asn Glu Lys Phe
50 55 60
 Lys Ser Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
65 70 75 80
 Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
85 90 95
 Ala Arg Tyr Tyr Tyr Gly Gly Ser Ser Xaa Xaa Val Tyr Xaa Tyr Trp
100 105 110
 Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
115 120 125

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 129 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
 Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
20 25 30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45
Gly Trp Ile Asn Pro Tyr Gly Asn Gly Asp Thr Asn Tyr Ala Gln Lys
50 55 60
Phe Gln Gly Arg Val Thr Ile Thr Ala Asp Thr Ser Thr Ser Thr Ala
65 70 75 80
Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr
85 90 95
Cys Ala Arg Ala Pro Gly Tyr Gly Ser Gly Gly Gly Cys Tyr Arg Gly
100 105 110
Asp Tyr Xaa Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125
Ser

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Gly Ser Gly Tyr Thr Phe Thr Ser Tyr
20 25 30
Trp Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
35 40 45
Gly Glu Ile Asp Pro Ser Glu Ser Asn Thr Asn Tyr Asn Gln Lys Phe
50 55 60
Lys Gly Arg Val Thr Leu Thr Val Asp Ile Ser Ala Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Gly Gly Tyr Asp Gly Trp Asp Tyr Ala Ile Asp Tyr Trp Gly
100 105 110
Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 30
- (D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CCCAAGCTTC CAGGGRCCAR KGGATARACN GRTGG

35

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CCCAAGCTTA CGAGGGGGAA GACATTGGG AA

32

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GGGAATTCAT GRAATGSASC TGGGTYWTYC TCTT

34

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ACTAGTCGAC ATGAAGWTGT GGBTRAACTG GRT

33

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CCCAAGCTTA CTGGATGGTG GGAAGATGGA

30

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ACTAGTCGAC ATGGATTTC ARGTCAGAT TWCAGCTT

39

(2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GGAAGCTTCC ACCATGGATT TCGGACTGGC CC

32

(2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GGACTAGTGG TTTGGACGAG CCTGTTG

27

146
concl